第二章

随机变量及其分布

第二章

2.1 随机变量：将样本映射为实数

$\begin{cases}\text{离散型随机变量：有限个，无限可列个}\\ \text{非离散型随机变量：连续型}\end{cases}$

2.2 离散型随机变量及其分布律

分布律：X 的所有值 x_k $(k=1,2\cdots)$ 可列个.

$$P\{X=x_k\}=p_k \quad (k=1,2\cdots)$$

三种重要的离散型随机变量

(一) 0.1 分布（特殊的二项分布）

(二) 二项分布：只有两个结果 A 及 $\bar{A}$，重复，独立 n 重伯努利

$X\sim B(n,p)$ $P\{X=k\}=C_n^k p^k(1-p)^{n-k}$ $k=0,1,2\cdots n$

最可能值：$(n+1)p$ 不为整 $[(n+1)p]$（取整）得最可能值

$(n+1)p$ 为整 $(n+1)p-1$ 为最可能值

(三) 泊松分布 $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$ $k=0,1,2,3\cdots$

$X\sim P(\lambda)$ $(\lambda>0)$（等车，收银台，电话呼叫，公用设施）

计算 → 查表

泊松分布逼近二项分布（n 较大，p 较小，np 适中）$\left(\begin{matrix}n\geq 100\\ np\leq 10\end{matrix}\right)$

$$C_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k e^{-\lambda}}{k!}$$

(四) 几何分布 $P(A)=p$，第 k 次首次发生，前 k-1 未发生

$P\{X=k\}=(1-p)^{k-1}p$ $X\sim G(p)$. $(k=1.2.3\cdots)$

(五) 超几何分布 N 个元素：N_1 个属一类 N_2 个属另一类

$P\{X=k\}=\frac{C_{N_1}^k C_{N_2}^{n-k}}{C_N^n}$ 取 n 个 X：n 个属于第一类的个数

No.

Date

2.3 随机变量的分布函数

1. 分布函数 $F(x)=P(X\leq x)$ $x\in(-\infty,+\infty)$ $F(x)\in[0,1]$

$P\{x_1<X\leq x_2\}=P\{X\leq x_2\}-P\{X\leq x_1\}=F(x_2)-F(x_1)$

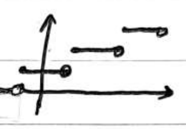
1°) $F(x)$ 不减 2°) $0\leq F(x)\leq 1$ 3°) $F(x+0)=F(x)$（右连续）

4°) $\lim\limits_{x\to+\infty}F(x)=F(+\infty)=1$ $\lim\limits_{x\to-\infty}F(x)=F(-\infty)=0$

5°) $P\{X=a\}=F(a)-F(a-0)$ 6°) $P\{a\leq X\leq b\}=F(b)-F(a-0)$

7°) $P\{X<a\}=F(a-0)$

2、离散型的分布函数



X	-2	0	1	3
P	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

$F(x)\begin{cases}0, & x<-2\\ \frac{1}{2} & -2\leq x<0\\ \frac{1}{2}+\frac{1}{4} & 0\leq x<1\\ \frac{1}{2}+\frac{1}{4}+\frac{1}{8} & 1\leq x<3\\ 1 & x\geq 3\end{cases}$ $P\{X=x_k\}=F(x_k)-F(x_k-0)$

3、连续型的分布函数 $F(x)=P(X\leq x)=\int_{-\infty}^{x}f(t)dt$

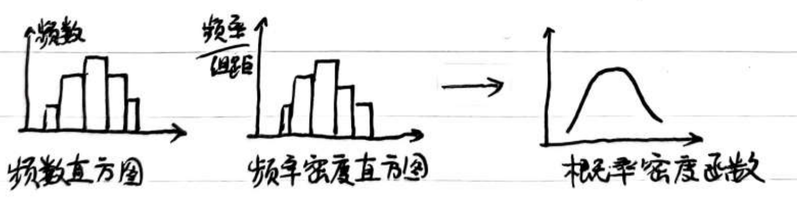
$f(x)=\begin{cases}-\frac{1}{2}x+1 & 0\leq x\leq 2\\ 0 & \text{其他}\end{cases}$

$x<0$ $F(x)=\int_{-\infty}^{x}0dt=0$

$0\leq x<2$ $F(x)=\int_{-\infty}^{x}f(t)dt=\int_{-\infty}^{0}\cdots+\int_{0}^{x}\cdots$

$x\geq 2$ $F(x)=\cdots=\int_{-\infty}^{0}\cdots+\int_{0}^{2}+\int_{2}^{x}\cdots$

2.4 连续型随机变量及其概率密度 (2.2 ~ 2.4)



$F(x)=\int_{-\infty}^{x}f(t)dt$ X：连续型随机变量

$f(x)$ 称 X 的概率密度

$P\{a<X\leq b\}=\int_a^b f(x)dx$

$X\sim f(x)$

KOKUYO

Date

性质：1°) $\int_{-\infty}^{+\infty}f(x)=1$ 2°) 连续变量取个别值的概率为 0

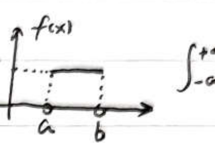
3°) 连续型端点无所谓（概率为 0 的事件未必不可能事件）

（概率为 1 的事件未必是必然事件）

三种重要的连续型随机变量

(一) 均匀分布 $X\sim U[a,b]$

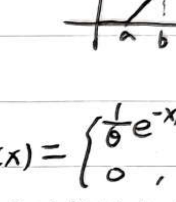
$f(x)=\begin{cases}\frac{1}{b-a}, & a\leq x\leq b\\ 0, & \text{else}\end{cases}$



$\int_{-\infty}^{+\infty}f(x)dx=1$

分布函数

$F(x)=\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & a\leq x<b\\ 1 & x\geq b\end{cases}$

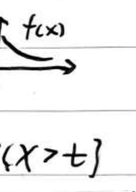


(二) 指数分布

$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\leq 0\end{cases}$ 或 $f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta}, & x>0\\ 0, & x\leq 0\end{cases}$

$X\sim Exp(\lambda)$ θ/λ 为常数，X 服从参数 θ 的指数分布

$F(x)=\begin{cases}1-e^{-x/\theta}, & x>0\\ 0, & \text{其他}\end{cases}$



无记忆性 $P\{X>s+t\mid X>s\}=P\{X>t\}$

一元件已使用 S 年后再可用七年的概率 = 全新时可用七年的概率

(三) 正态分布 $X\sim N(\mu,\sigma^2)$

$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty<x<+\infty$

对称轴 $x=\mu$ $\varphi(x)_{max}=\frac{1}{\sqrt{2\pi}\sigma}$ μ：左右移 σ：σ 变小，最高点上移

标准正态分布 $\mu=0$ $\sigma=1$

$\varphi_0(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ $\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$

$(\mu-\sigma,\mu+\sigma)$ 68.26%

$(\mu-2\sigma,\mu+2\sigma)$ 95.44%

$(\mu-3\sigma,\mu+3\sigma)$ 99.74%

$\varphi_0(x)=\varphi_0(-x)$ $\Phi_0(-x)=1-\Phi_0(x)$（分布）

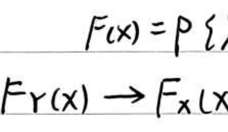
$\varphi(x)=\frac{1}{\sigma}\varphi_0\left(\frac{x-\mu}{\sigma}\right)$ $\Phi(x)=\Phi_0\left(\frac{x-\mu}{\sigma}\right)$

一般 → 标准

KOKUYO

Date

上 α 分位数



$P\{X>z_\alpha\}=\alpha$, $0<\alpha<1$

α	0.001	0.005	0.01	0.025	0.05	0.10
z_α	3.090	2.576	2.326	1.960	1.645	1.282

2.5 随机变量函数的分布

已知随机变量，构造函数，问其分布

1°) 离散型

X	-1	0	1	2
p	0.2	0.3	0.4	0.1

$\xrightarrow{Y=X^2}$

Y	0	1	4
p	0.3	0.6	0.1

2°) 连续型 $F(x)=P\{X\leq x\}$

1. 由 $F_Y(x)\to F_X(x)$ 2. 两边求导 3. $f_X(x)\to f_Y(x)$

例：$f_X(x)$ $Y=3X+2$

解 $F_Y(x)=P\{Y\leq x\}=P\{3X+2\leq x\}=F_X\left(\frac{x-2}{3}\right)$ 求导

$f_Y(x)=\frac{1}{3}f_X\left(\frac{x-2}{3}\right)$

特别的：X 服从 $[a,b]$ 均匀分布，$Y=kX+C$ 服从相应区间均匀分布

$f_X(x)=\begin{cases}\frac{1}{b-a}, & a\leq x\leq b\\ 0, & \text{else}\end{cases}$ $f_Y(x)=\begin{cases}\frac{1}{kb-ka}, & ka+c\leq x\leq kb+c\\ 0, & \text{else}\end{cases}$

正态分布 $X\sim N(\mu,\sigma^2)$ $Y=aX+b$

$Y\sim N(a\mu+b, a^2\sigma^2)$

定理：X 的密度 $f_X(x)$ $Y=kX+b$ $(k\neq 0)$，$f_Y(x)=\frac{1}{|k|}f_X\left(\frac{x-b}{k}\right)$

本笔记在https://github.com/dydcyy-gh/study-notes开源